

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12P 19/34, A61K 38/00	A1	(11) International Publication Number: WO 98/11249 (43) International Publication Date: 19 March 1998 (19.03.98)
(21) International Application Number: PCT/US96/14708 (22) International Filing Date: 13 September 1996 (13.09.96) (71)(72) Applicant and Inventor: GARVIN, Alex, M. [US/FR]; 13, rue du Chemin de Fer, F-68480 Durnenach (FR).		(81) Designated States: US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: MUTATION DETECTION USING PEPTIDE TAGGED <i>IN VITRO</i> SYNTHESIZED PROTEINS (57) Abstract A process for detecting protein altering mutations in genes. Coding sequence of the gene is PCR amplified with a 5' primer that contains at its 5' end a polymerase binding site, a translation initiation site, and an in frame sequence coding for a peptide tag, followed by the in frame 5' end of the test sequence. After PCR amplification of the test sequence, the PCR product is used as template to make mRNA in an <i>in vitro</i> transcription reaction using an RNA polymerase that recognizes the polymerase binding site incorporated into the 5' PCR primer. The mRNA is then used as template to make protein in an <i>in vitro</i> translation reaction. The protein encoded by the test sequence has at its amino terminus a peptide tag that can be used to either detect the protein, or to purify the protein for further analysis.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Mutation Detection Using Peptide Tagged *In-Vitro* Synthesized Proteins

Background- Field of Invention

This invention relates to a process for detecting mutations in genes by synthesizing the proteins encoded by the genes *in-vitro* with a peptide tag attached to the amino termini, and analyzing properties of the tagged proteins.

Background- Description of Prior Art

Many genes that predispose to human disease have now been identified and hundreds more will be found over the next decade. In order to take advantage of these findings, the biomedical community must be able to identify individuals that carry mutations in these genes in order to counsel them and to help them take preventative measures.

The methods now being used to screen for mutations in certain disease causing genes suffer from short comings that will make it far too time consuming and expensive to carry out large scale screenings. For example, the breast cancer susceptibility genes BRCA1 (1) and BRCA2 (2) have recently been identified and mutations predisposing to breast cancer have been found throughout the 16 kilobases of coding sequence contained in these two genes. The entire coding sequence of

BRCA1 and BRCA2 are now routinely screened by either DNA sequencing (3), Single Strand Conformation Polymorphism (SSCP) (3), or the Protein Truncation Test (PTT) (4). All three of these techniques rely on the electrophoretic properties of either DNA or protein, and detection of the electrophoresed products is accomplished by incorporation of either a fluorescent or radiolabeled tag. Sequencing both strands of 16 kb for a single individual is tedious with current techniques and gives a large amount of unnecessary information, since virtually all of the sequence will be identical to the known sequence. SSCP is less tedious, but still requires sequencing benign polymorphisms. Furthermore SSCP segment length is limited to 150 base pairs, requiring a minimum of 6 PCR reactions per kilobase screened. PTT is the fastest of these techniques but it is incapable of detecting missense mutations, and it also suffers from all of the drawbacks of gel electrophoresis, such as gel preparation, long running times, and low resolution. In addition, PTT will miss some mutations due to artifactual proteins smaller than the wildtype protein that can obscure the signal from a bonafide truncation product. These artifacts can be due to ATG codons within the coding sequence that act as cryptic translation initiation sites. Proteins of this type lack proper amino termini, and hence would lack a peptide tag that one might engineer into the amino termini.

Other techniques such as Allele Specific Oligomer Hybridization (5) and Allele Specific PCR (6) are designed to identify one previously characterized mutation per assay. These techniques are useful when a particular mutation has already been identified and is widespread, but cannot be used for genes where a large number of yet unidentified mutations exist in the human population.

Objects and Advantages

The objects and advantages of my patent are as follows:

- a) to provide a method capable of detecting mutations that qualitatively alter the protein product of any protein encoding gene, using genomic DNA or cDNA as starting material;
- b) to provide a method for mutation detection that can take advantage of the speed, high throughput, sensitivity, and accuracy of measuring devices that require purified materials for analysis (such as electrospray mass spectrometry);
- c) to improve the protein truncation test by eliminating artifacts caused by cryptic translation initiation.

Figures

Fig 1 shows the features present in the 5' primer used to PCR amplify a region of coding sequence to be tested.

Fig 2 shows a flow chart of how the process works. The 6 major steps are numbered and are discussed below.

Description of Figures

Figure 1

The 5' primer shown in fig 1 is a standard synthetic oligonucleotide without any modifications at either the 5' or 3' ends. The unique features of the oligonucleotide lie in the functional properties of its sequence. Going from the 5' to 3' direction, the following 4 elements are present in the order given:

- 1) A sequence recognized by an RNA polymerase. An example is the sequence recognized by the viral T7 polymerase TAATACGACTCA-CTATAGGG.
- 2) A sequence that allows translation initiation of mRNA. An example is the methionine codon placed after the Kozak consensus: AGACCACCATG.
- 3) An in frame sequence coding for a peptide that can be used for detection or affinity purification (a peptide tag). An example is the sequence GAC TAC AAG GAC GAC GAT GAC AAG, which codes for the FLAG peptide DYKDDDDK. When the FLAG peptide is present at the amino terminus of a protein, it is recognized by the anti-FLAG M2 monoclonal antibody. The M2 antibody can be used in western blotting to detect FLAG containing proteins, or it can be used to affinity purify the protein for further analysis.
- 4) An in frame sequence complimentary to the 5' end of the coding sequence being tested (the test sequence). This 5' end of the test sequence must be of sufficient length to allow the entire oligomer to hybridize to the test sequence present in the genomic DNA or cDNA sample and to act as a primer for PCR. Usually 20 bases is enough. An example is the in frame sequence GCT TGT GAA TTT TCT GAG AC, which is complimentary to the 5' end of exon 11 of the BRCA1 gene.

Figure 2

The flow chart in fig 2 shows how the whole process works.

1) Preparation of template for PCR. The PCR reaction must amplify a continuous stretch of coding sequence. The maximum length of coding sequence that can be analyzed per segment is certainly greater than 2 kilobases, thus for genes having many small exons it will be convenient to use cDNA as template. If the mutant message is unstable, this will present a problem, because the absence of the mutant message will result in the lack of mutant cDNA. When the mutant message is unstable in diploid organisms the only template in cDNA derived from a heterozygous carrier would be the wildtype. My experience with the BRCA1 gene has been that the problem of "non-sense mediated mRNA decay" (7), in which message instability is caused by a mutation that generates a premature stop codon, is not really a problem, and I expect that for most genes cDNA will be the template of choice.

The source of the tissue used to extract RNA must express the gene being tested. Many human genes (if not most) are expressed to some extent in peripheral blood, and can be PCR amplified from cDNA derived from peripheral blood leukocytes. Genomic DNA can be conveniently used as template for the coding sequence of large exons. It can also be used when cDNA is not available, and when mutant message is thought to be unstable. It should be noted that splice site mutations within introns and a subset of those within exons will be not be detected when genomic DNA is used as template.

2) PCR. A continuous region of coding sequence for the gene of interest is amplified by the polymerase chain reaction using the 5' primer described in fig 1, and a 3' primer that need not be in-frame. When cDNA is used as template and the gene is expressed at low levels in the source tissue used to make RNA, nested PCR is often required to get good amplification. An aliquot of the PCR product is used for the next step.

3) *In-Vitro* Transcription. RNA is generated by using the PCR product from step 2 as template in a reaction mixture containing all of the other components required for RNA synthesis. The RNA polymerase used in

this step must recognize the promoter sequence at the 5' end of the PCR product.

4) *In-Vitro* Translation. The RNA made in step 3 is then used as template in a cell lysate containing all the reagents required for protein synthesis. Rabbit reticulocytes and wheat germ are common lysate sources. Steps 3 and 4 could be done simultaneously using a coupled transcription/translation system, but this is not essential. Radiolabelled amino acids are included in the lysate if autoradiography will be performed later.

5a) Analysis of Unpurified Protein Using Tag. The unpurified tagged protein made in step 4 can be analyzed directly using ligands specific for the peptide tag. For example, the size of the protein can be determined by western blotting using an antibody directed against the tag. Note that artifactually small proteins resulting from cryptic translation initiation will not be detected by ligands specific for the amino terminal tag.

5b) Affinity Purification. The protein made in step 4 is present in a lysate containing thousands of different proteins, ribosomal RNA, salts, metabolites, lipids, etc. In order to analyze this protein using some of the more sensitive techniques available (such as electrospray mass spectrometry), it is necessary to purify the protein. The peptide tag engineered into the amino-terminus of the protein can be used for this purpose. One way to accomplish the purification is to use a ligand specific for the peptide tag immobilized to a solid surface. For example, if the FLAG peptide is used as the peptide tag, an anti-FLAG antibody could be covalently attached to magnetic beads. After incubation of the lysate with the anti-FLAG beads to allow binding, the beads could be washed, and the test protein could be eluted from the beads in a suitable buffer.

6) Analysis of the Purified Protein. The purified protein can now be analyzed by a variety of techniques that measure the physical, electrical, chemical, or biological properties of the protein. These properties are compared to the properties of the protein encoded by the wildtype sequence and any differences between the test protein and the control indicate that a mutation may be present. The assay used at this stage must be sensitive enough to analyze the amount of material purified from the lysate. For example, 50 microliters of a rabbit reticulocyte lysate may provide 10 picomoles of a 90 kD protein

and autoradiography of radiolabelled protein is certainly sensitive enough to detect this amount of protein. Such an approach will be similar to the standard PTT except that the background signals caused by labelled artifactual proteins lacking the tag will be absent. Another method of analysis is electrospray mass spectroscopy which can measure the mass of a protein to within 0.01% when sub-picomole amounts of the protein are available. This level of accuracy should allow for detection of mutant proteins having amino acid substitutions as well as truncations.

Summary, Ramifications, and Scope

In conclusion, the reader will see that tagged *in-vitro* synthesized proteins have advantages over untagged proteins in terms of how they can be analyzed and the quality of data obtained. Detection of the tag, rather than of all newly synthesized proteins from the lysate, eliminates the background noise caused by artifactual proteins lacking the tag. Furthermore, the tag can be used to purify the protein away from other components of the lysate, including the artifactual proteins lacking the tag, prior to analysis. To date, the exquisite accuracy, sensitivity, and throughput of certain protein analytical techniques have yet to be applied to mutation detection. The ability to easily purify *in-vitro* synthesized proteins may facilitate the use of these techniques in helping to identify disease causing genes in humans. Although I have only used the examples of autoradiography and mass spectroscopy as methods for analyzing the purified proteins, this should not be construed as limiting the scope of the invention, since other techniques may also be used. Rather, the scope of the invention should be determined by the appended claims.

References

- 1) Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994;266(5182):66-71.
- 2) Wooster R, Bignell G, Lancaster J, Swift S, Seal S, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;378:789-792.

- 3) Shattuck-Eidens D, McClure M, Simard J, et al. A collaborative survey of 80 mutations in the BRCA1 breast and ovarian cancer susceptibility gene: Implications for presymptomatic testing and screening. *Jama* 1995;273(7):535-541.
- 4) Hogervorst FBL, Cornelis RS, Bout M, et al. Rapid detection of BRCA1 mutations by the protein truncation test. *Nature genetics* 1995;10:208-213.
- 5) Struwing JP, Abeliovich D, Peretz T, et al. The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nat Genet* 1995;11(2):198-200.
- 6) Fitzgerald MG, Macdonald DJ, Krainer M, et al. Germline BRCA1 mutations in jewish and non-jewish women with early-onset breast cancer. *N Engl J Med* 1996;334:143-149.
- 7) Zhang ZX, Wakamatsu N, Mules EH, Thomas GH, Gravel RA. Impact of premature stop codons on mRNA levels in infantile Sandhoff disease. *Hum Mol Genet* 1994;3(1):139-45.

Claims: I claim:

- 1) A method for detecting protein altering mutations in genes which relies on analysis of *in-vitro* synthesized proteins encoded by said genes where said proteins have been tagged at their amino termini with a peptide sequence and where ligands recognizing said peptide sequence are used to purify said proteins for further analysis.
- 2) The analysis of claim 1 wherein said analysis is performed using electrospray mass spectroscopy.
- 3) The analysis of claim 1 wherein said analysis is performed using SDS/PAGE-autoradiography.
- 4) A method for detecting protein altering mutations in genes which relies on analysis of *in-vitro* synthesized proteins encoded by said genes where said proteins have been tagged at their amino termini with a peptide sequence and where ligands recognizing said peptide sequence are used to detect said proteins.
- 5) The detection of claim 4 wherein said detection is performed by western blotting.

Note to the Examiner: I choose not to claim the 5' oligomer shown in figure 1 because it is obvious once the concepts presented above are considered.

Figure 1 5' Primer for PCR

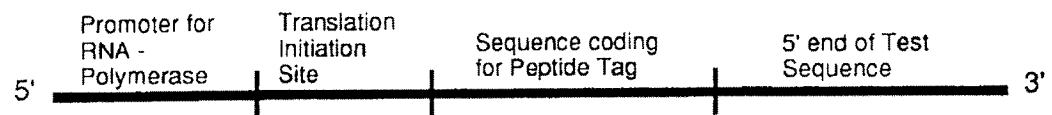
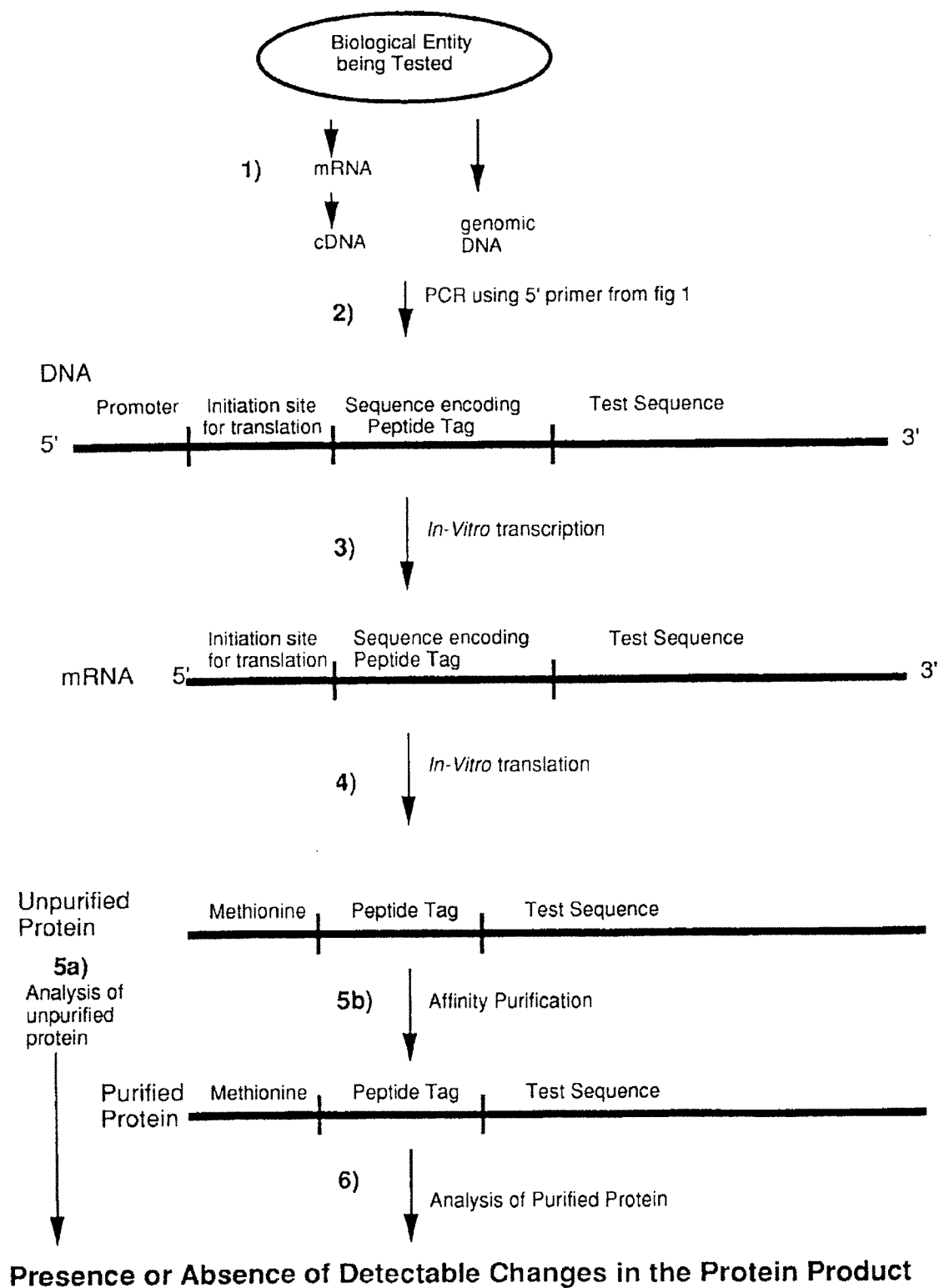


Figure 2 Flow Chart of the Process



INTERNATIONAL SEARCH REPORT

Intern. application No.

PCT/US96/14708

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12P 19/34; A61K 38/00

US CL : 435/91.2; 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/91.2; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

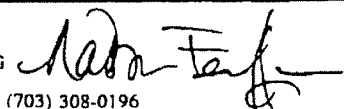
STN, USPAT MEDLINE, BIOSIS, CAPLUS
search term: pcr, primers, translate, tagged

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	HOGERVORST et al. Rapid Detection of <i>BRCA1</i> Mutations by the Protein Truncation Test. Nature Genetics. June 1995, Vol. 10, pages 208-212, especially pages 209, 212.	1, 4 ----- 2, 3, 5
Y	POWELL et al. Molecular Diagnosis of Familial Adenomatous Polyposis. New England Journal of Medicine. 30 December 1993, Vol. 329, No. 27, pages 1982-1987, especially pages 1983-1984.	2, 3, 5

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 25 NOVEMBER 1996	Date of mailing of the international search report 12 DEC 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer JOYCE TUNG  Telephone No. (703) 308-0196